基于核密度估计的基本概率指派生成方法 *

黄 杰 [1,2]，尉永清 [3†]，伊 静 [1,4]，刘孟迪 [1,2]

(1. 山东师范大学 信息科学与工程学院，济南 250358；2. 山东省分布式计算机软件新技术重点实验室，济南 250014；3. 山东警察学院 公共基础部，济南 250014；4. 山东建筑大学 计算机科学与技术学院，济南 250014)

摘　要： D-S 证据理论是一种有效处理不确定信息的方法，被广泛应用于各领域中，而 D-S 合成方法作用的对象是基本概率指派(basic probability assign,BPA)，如何生成 BPA 仍是 D-S 理论应用中重要且有待解决的首要步骤。针对生成 BPA 提出一种基于核密度估计 KDE（kernel density estimation）的 BPA 生成方法：训练数据用于构建基于最优化窗宽的核密度估计的数据属性模型；然后利用训练数据的核密度模型计算测试数据的密度—距离—分布值 Tri-D(density-distance-distribution)，通过嵌套式的方法分配 Tri-D 值获取测试数据对应的 BPA；最后 D-S 合成 BPA 得到最终判断，通过分类准确率来判断 BPA 生成方法的有效性。实验通过在 UCI 数据集上的与其他方法的分类准确率对比验证了提出方法的有效性。

关键词： 基本概率指派 BPA；核密度估计；Tri-D；窗宽

中图分类号： TP391　　doi: 10.19734/j.issn.1001-3695.2018.11.0882

New method to determine basic probability assign based on kernel density estimation

Huang Jie[1, 2], Wei Yongqing[3 †], Yi Jing[1, 4], Liu Mengdi[1, 2]

(1. *School of Information Science & Engineering*, *Shandong Normal University*, *Jinan* 250358, *China*; 2. *Shandong Provincial Key Laboratory for Distributed Computer Software Novel Technology*, *Jinan* 250014, *China*; 3. *Basic Education Dept*, *Shandong Police College*, *Jinan* 250014, *China*; 4. *School of Computer Science & Technology*, *Shandong Jianzhu University*, *Jinan* 250014, *China*)

Abstract: D-S evidence theory is a method that processes uncertain information effectively and is widely used in information fusion. However, the determination of BPA (Basic Probability Assign) or the action object of D-S fusion method is still an open problem in process of D-S theory application. This paper proposed a BPA determination method based on kernel density estimation(kde) . The method uses training data to construct a data attribute model with optimized bandwidth based on the optimized kernel density estimation; then calculate the density-distance-distribution (Tri-D) value of test data by using the kernel density model of training data. The next step is obtaining BPA of test data by using the nested method to assign Tri-D. Finally, fusing BPA by D-S method to get the final result, and judging the validity of the BPA generation method by the classification accuracy rate. An illustrative case regarding the classification accuracy compared with other methods on UCI data sets shows the effectiveness of the method.

Key words: basic probability assign(BPA); kernel density estimation; Tri-D; bandwidth

0　引言

D-S 证据理论是由 Dempster[1]提出，并由 Shafer[2]发展的一种有效处理不确定信息和组合多元信息的方法。它和经典的贝叶斯理论是数据融合中两个主流框架，但相比于贝叶斯理论，D-S 理论具有将概率同时分配给单子集和复合子集的优点。D-S 理论在数据融合[3]、综合评估[4,5]、分类[6,7]、进化博弈论[8~10]等诸多领域实现了很好地应用。在应用 D-S 框架的过程中，BPA 的生成是关键又核心的第一步，会对最终的结果产生很大的影响。但如何生成 BPA 到现在依旧没有通用的解决办法，不少学者就此进行了研究，并提出了不同的解决方法。较早的 Yager[11]引进了与 D-S 信念结构相关的一整类模糊测度，并讨论了模糊测度的熵；蒋雯等人[12~14]提出了基于广义模糊数生成 BPA 的方法；Liu 等人[15]将模糊朴素贝叶斯与最近邻均值分类进行加权结合提出了 WFDSF 算法；文献[16]提出了一种利用训练数据核心样例获得 BPA 的方法；文献[17]提出了基于分类器含混矩阵的方法；此外，还有基于聚类确定 BPA 的方法[18,19]。

通过对以上文献方法的学习可以发现，BPA 的生成本质上是各可能事件发生概率的确定，而核密度估计正是一种完全基于数据本身来进行概率密度估计的有效非参数估计方法，所以本文提出一种基于核密度估计的 BPA 生成方法，首先利用训练数据构建最优化窗宽（以下简称 h）的核密度估计模型；计算测试数据各属性在各模型中核密度取值，并进一步计算 *Tri-D*，通过嵌套式分配[20]生成 BPA；D-S 合成 BPA 得到判定结果，结合 UCI 数据集进行本文方法有效性的验证。

收稿日期： 2018-11-17；**修回日期：** 2019-01-21　**基金项目：** 国家自然科学基金资助项目(61373148)；山东省自然科学基金资助项目(ZR2014FL010)；山东省教育厅基金资助项目（J15LN34）；山东省社科规划项目（17CHLJ18，17CHLJ33，17CHLJ30）

作者简介： 黄杰（1992-），女，山东济南人，硕士研究生，主要研究方向为数据挖掘；尉永清（1963-），女（通信作者），教授，硕导，主要研究方向为网络信息安全（13864155372@163.com）；伊静（1979-），女，副教授，博士研究生，主要研究方向为数据挖掘；刘孟迪（1993-），女，硕士研究生，主要研究方向为数据挖掘.

1 相关工作

1.1 D-S 证据理论

下面对 D-S 证据理论的基本概念进行介绍。

a) 辨识框架(frame of discernment)。

D-S 理论中，将包含所有可能发生假设的集合 $\Theta=\{H_1,H_2,...,H_n\}$ 定义为辨识框架，其中的假设是穷尽且独立的，Θ 的幂集为 $\Omega=\{\{\varphi\},\{H_1\},...,\{H_n\},\{H_1H_2\},...,\{\Theta\}\}$，共 2^n 个假设集合。

b) 基本概率指派(basic probability assign,BPA)。

设幂集 Ω 到[0,1]的映射 m 满足式(1)，则称 m 为基本概率指派（下文简称 BPA），又称 mass 函数或证据，其中 m 值非零的子集称为焦元。

$$\begin{cases}\sum_{A\in\Omega}m(A)=1\\ m(\varphi)=0\end{cases} \tag{1}$$

m 值表示对子集的支持程度，值越大表示支持程度越大。

c) 组合规则。

对于两条证据，D-S 组合规则如下，其中 B_i 和 C_j 为焦元，k 称为冲突系数。

$$\begin{cases}m(A)=\frac{1}{1-k}\sum_{B_i\cap C_j=A}m_1(B_i)m_2(C_j),(A\neq\varphi,A\subset\Theta)\\ m(\varphi)=0\end{cases} \tag{2}$$

$$k=\sum_{B_i\cap C_j=\varphi}m_1(B_i)m_2(C_j) \tag{3}$$

1.2 Pignistic 概率[21]

在 D-S 组合完 BPA 之后，取最大 Pignistic 概率值为最终的结果，计算公式如式（4）所示 ，其中|X|表示集合 X 的基数。

$$P_{pig}(A)=\sum_{B\subseteq\Theta}\frac{|A\cap B|\times m(B)}{|B|} \tag{4}$$

介绍完 D-S 证据理论基本内容，对 D-S 应用与本文工作的关系进行简单说明。本文工作是将有不同特征的数据转换为形如 m(A)=0.3,m(B)=0.6,m(C)=0.1 且满足式(1)的 BPA(A,B,C 就是 Θ 中的假设，也就是数据可能的类别或标签)以供 D-S 按式(3)(4)进行组合，得到最终一条 BPA，所以 BPA 生成是 D-S 应用过程必不可少的一步。

1.3 核密度估计(KDE)

求解给定样本集合分布密度的方法包括参数估计和非参数估计两种类型。核密度估计（或称 Parzen 窗法）属于后者。与参数估计中需要假定所估计数据在各个可能类别中都服从特定分布的“不准确”前提相比，由 Rosenblatt[22]和 Parzen[23]提出的核密度估计从数据本身出发研究数据的分布特征，在统计和各应用领域均受到了高度重视。对于独立同分布随机变量 $x_1,x_2,...,x_n$,其真实服从的概率密度函数 f(x)的核密度估计函数表示如下：

$$\hat{f}_h(x)=\frac{1}{nh}\sum_{i=1}^{n}K(\frac{x_i-x}{h}),x\in R \tag{5}$$

其中：$K(\bullet)$ 为核函数；h 为预先给定的正数，通常称为窗宽或光滑参数。为方便起见，本文记 $K_h(\mu)=K(\mu/h)/h$，则式(5)可以表示为

$$\hat{f}_h(x)=\frac{1}{n}\sum_{i=1}^{n}K_h(x_i-x),x\in R \tag{6}$$

从式(6)可以看出，f 的密度估计 $\hat{f}$ 不仅与给定的样本集合有关，还与核函数的选择与带宽参数的选择有关。常见的核函数包括高斯核函数(normal)、三角核函数(triangle)、均匀核函数(uniform)和 Epanechnikov 核函数等,本文在此不再赘述。本文主要考虑窗宽 h 因素，h 的取值决定了密度曲线的光滑程度。图 1 是一组随机生成的服从正态分布的数据在不同 h 时的核密度估计曲线。

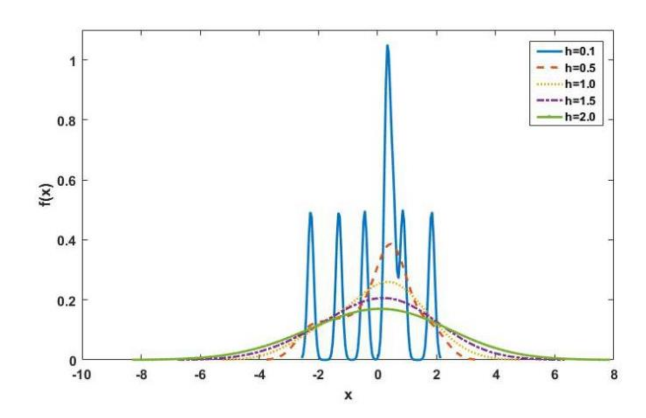


图 1 正态分布数据在不同 h 下的 kde 曲线

Fig. 1 KDE curves of random data with different h

由图 1 可以看出，当 h 取值过小时，曲线过于陡峭波折，出现了过拟合现象；而随着 h 的增大，曲线逐渐平缓；当 h 增大到 2 时，曲线就过于平滑，使数据的特征被掩盖。综上，选择合适的 h 非常重要，本文采用最优化的方法来选择 h，详见 2.1 节。

2 基于核密度估计的基本概率指派生成方法

BPA 生成的基本要求是要满足后续组合过程的使用，在此基础上要尽可能多地提取和利用数据信息，不同的属性可以从不同的角度反映数据的特征，因此本文以属性为基本单位进行 h 的优化和 kde 属性模型的构建，后续的 BPA 分配也是基于属性模型的（需要说明的是，将一维属性下得到的关于 v 个类别的 v 个 kde 子曲线统称为一个属性模型）。

下面将介绍本文的方法，首先给出以下假设：所有可能的假设或最终所有的类别共 v 个，即辨识框架表示为 $\Theta=\{c_1,c_2,...,c_v\}$，数据集包含 n 条数据，每条数据有 p 维属性，x_t 表示测试数据。

2.1 基于最优化窗宽的属性 kde 模型

现有窗宽的优化方法主要是基于积分均方误差 MISE(mean integrated squared error)的优化[24],MISE 的定义如下：

$$MISE(h)=E[\int(\hat{f}_h(x)-f(x))^2dx]=\int[E\hat{f}(x)-f(x)]^2dx+\int Var\hat{f}(x)dx$$

上式中分解后的第一项表示期望值与真实值间的偏差，第二项表示估计值的方差，对上式求解可以得到偏差及方差的表示式，经过求导、求最小值等一系列过程最终推导出最优窗宽 h 的表达式如下(详细公式推导见文献[24])：

$$h=(\frac{4\hat{\sigma}^5}{3n})^{0.2}\approx1.059\hat{\sigma}n^{-0.2} \tag{7}$$

其中：$\hat{\sigma}$ 和 n 分别代表样本的标准差和数目。根据式(7)可以计算出训练数据属性 j 对应类别 i 的最优窗宽 h_{ij}^*,最终得到训练数据的 p 行 v 列的最优化窗宽矩阵 $\mathbf{H}^*$。

利用训练数据构建属性核密度估计模型，具体操作为：对训练数据属性 j 按不同类别进行划分，不同属性下不同类

别的优化窗宽 h 在 $\mathbf{H}^*$中已经相应求出，按式(6)进行核密度估计，最终一维属性下共包含 v 个子 KDE 模型，按照之前第 2 章开始的说明，本文称 p 维属性共对应 p 个模型。

2.2　密度—距离—分布值(Tri-D)

基于训练数据的 kde 属性模型构建已完成，测试数据的属性在各属性模型下的密度取值可随之得出，数据在某类别下的密度值的大小与数据属于该类别有正相关的关系，但在类间距离较近或密度差距较明显的情况下，仅根据密度值大小来判断所属类别可能会导致较多错误的发生。在区间[4,5]和[5.2,5.6]分别随机生成 10 个数据记为 A 类和 4 个数据记为 B 类，通过核密度估计得到它们的 kde 曲线如图 2 所示。

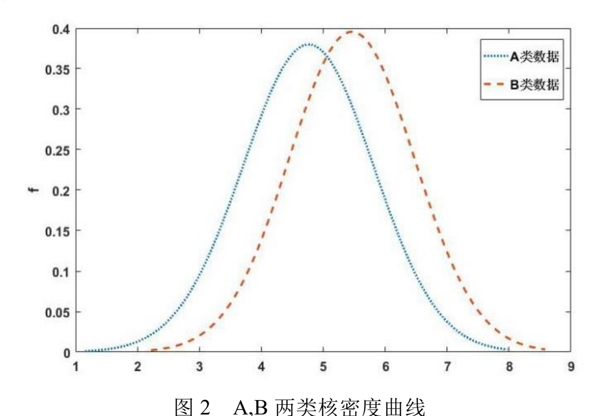


图 2　A,B 两类核密度曲线

Fig. 2　KDE curves of A and B classes

由图 2 可以发现类，别 A 与 B 的密度曲线在 x 轴重叠部分较大，可以想象随着 A,B 类间距离的缩小（在另一方不动的前提下，A 曲线往右移动或者 B 曲线往左移动），这将导致两曲线重叠部分越来越大，这样一来仅靠密度值已经无法正确区分点的类别。

在传统的分类聚类算法中,KNN 通过距离选择样例的 k 近邻进行分类；K-means 算法通过样例与中心点间的距离来进行聚类，所以距离是类别判断中非常重要的因素；另外，同一类别的点之间的距离分布应该相似度更高，所以结合密度值、距离以及分布特征，本文提出了 *Tri-D* 值的概念，测试点 x_t 的属性 j 在相应 j 属性模型中关于类别 i 的 *Tri-D* 值定义如下，由此对于有 v 个类别的数据来说，在一维属性下共可以得到 v 个 *Tri-D* 值：

$$Tri-D_{tij}=f_{ij}(x_t{}^j)\cdot(|d_{tij}\cdot\overline{d_{tj}}^{-1}-1|)^{-1} \tag{8}$$

其中：$f_{ij}(x_t{}^j)$ 代表点 x_t 的属性 j 在对应第 j 维属性模型中 i 类曲线上的密度取值；d_{tij} 表示 x_t 属性 j 的值到训练数据属性 j 的第 i 类数据中心点的距离；中心点矩阵 $Cen_{p\times v}$ 由算法 k-means++[25]对每维属性的各类别聚类得到；$\overline{d_{tj}}$ 代表训练数据的第 j 维属性下第 i 类数据中任意两点之间距离的均值。

对上述定义进行简单分析：首先第一部分 $f_{ij}(x_t^j)$，其取值越大说明在属性 j 的条件下 x_t 属于类别 i 的可能性越大；其次，第二部分由两个距离组成：第一个距离是测试点到训练数据类别中心点的距离，第二个距离选择了训练数据类别 i 中任意两点之间的平均距离；通过前面的两个距离的比较来实现分布特征的考量，如果两个距离越相近，即 $|d_{tij}\cdot\overline{d_{tj}}^{-1}-1|$ 越小，本文就认为 x_t 与该类别所有点的距离分布越相似。所以综合以上分析，$Tri-D_{tij}$ 值越大，在属性 j 的条件下 x_t 属于类别 i 的可能性越大；反之，则越小。

2.3　基本概率指派生成

每个测试数据 x_t 的属性 j 在相应属性模型下可以得到 v 个 *Tri-D* 值，所有 p 个属性共可以得到 p 行 v 列的 *Tri-D* 值矩阵。以 x_t 的任一维属性 j 为例说明本文的嵌套式[20]的分配方法:属性 j 在对应模型中可以得到 v 个 *Tri-D* 值组成向量 $Tri\text{-}D_{tj}$，对 $Tri\text{-}D_{tj}$ 按降序排序,得到，设相应的 $Tri-D_{tj}{}'$ 类别向量为 $\mathbf{C_d}$，则本文分配方式如下：

$$\begin{aligned}&m(C_d[1])=Tri-D_{tj}{}'[1]\\&m(C_d[1],C_d[2])=Tri-D_{tj}{}'[2]\\&\quad\cdots\cdots\\&m(C_d[1],C_d[2],\ldots,C_d[v])=Tri-D_{tj}{}'[v]\end{aligned} \tag{9}$$

通过这种分配方法可以同时得到单焦元 BPA 和多焦元复合 BPA。为满足式(1)的要求，需要对分配得到的各式进行归一化得到所需的 BPA。最终，x_t 的每一维属性都可以得到一条 BPA，p 维属性得到 p 条 BPA，通过组合得到最终 BPA。

2.4　算法流程

本文算法的流程给出如图 3 所示。下面对算法的相关细节进行介绍：将数据分为用来学习和构建属性模型的训练数据和评估方法有效性的测试数据，首先利用训练数据构建 kde 模型，将每一维属性都视为一个信息来源：对于每一维属性 j 中的不同类别 i，通过 3.1 中式(7)计算得到相应的窗宽 $h_{ij}{}^*$，最终得到最优窗宽矩阵 $\mathbf{H}^*$；对各属性对应的各类别数据按照相应的优化窗宽进行核密度估计(本文使用高斯核函数，见 4.2 节)，得到属性 j 的 kde 模型。对于测试数据：根据构建的训练数据属性模型，按式(8)计算 *Tri-D* 值，最终任一条测试数据 x_t 都对应得到值矩阵 $Tri-D_{p\times v}$，依照式(9)进行分配获得 p 条 BPA，D-S 组合得到最终 BPA，取最大 Pignistic 值对应假设为最后的结果。

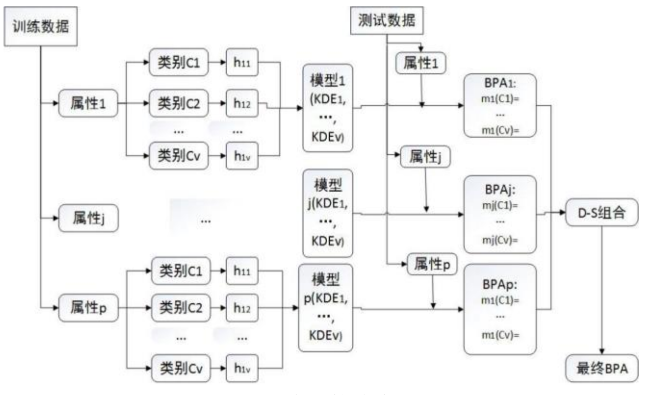


图 3　本文算法流程

Fig. 3　Procedures of proposed method

3　实验及结果分析

本节共有三组实验在 UCI 数据集上的实验：3.1 节对 Iris 数据集的数据进行实验流程的示例说明；3.2 节确定了核函数的选择，并对本文窗宽的优化效果进行了证明；3.3 节通过在 UCI 数据集上的分类准确率说明了本文方法的有效性。

本文实验中用到的 UCI 数据集的名称、实例数、类别数、属性数以及是否有缺失值等信息由表 5 给出。对于缺失值的情况在应用 D-S 证据理论的过程中可以直接忽略无须做处理，这也是 D-S 理论的优点之一。

3.1　本文方法实验示例

本节是通过对 Iris 一测试数据的分类过程来对本文算法流程进行说明。Iris 数据集是分类基准数据集，共有三个类别 Setosa(S)、Versicolor(C)、Virginica(V)。每条数据包含 SL,SW,PL,PW 四维属性，每个类别均有 50 个样例，共 150

个样例。本示例实验取 80%为训练数据，给定标签为 C 的测试样例,其各属性取值如下：SL=6.1cm, SW=2.9cm, PL=4.7cm, PW=1.4cm。

首先利用训练数据构建属性 kde 模型：利用式(7)计算训练数据各属性的最优窗宽，得到各属性中不同类别的最优窗宽，如表 1 所示。图 4~7 展示了通过核密度估计构造的各属性模型及模型中不同类别的核密度曲线。其次，计算测试数据的 *Tri-D* 值：利用 K-means++[25]聚类得到训练数据各属性不同类别的中心点，如表 2 所示，按照 *Tri-D* 的定义计算得到 *Tri-D* 矩阵，如表 3 所示；生成测试数据的 BPA：通过嵌套式分配方法得四维属性对应的共计四条 BPA，如表 4 所示，利用式(3)计算任意两 BPA 之间的冲突系数 k，结果为 $k_{12}=k_{13}=k_{14}=k_{23}=k_{24}=k_{34}=0$，BPA 之间不存在冲突问题；D-S 组合：组合得最终 BPA 为 $m(\mathrm{S})=m(\mathrm{V})=0, m(\mathrm{C})=1$，所以判定测试数据属于“最被支持”的 C 类，与真实类别一致，判定正确。

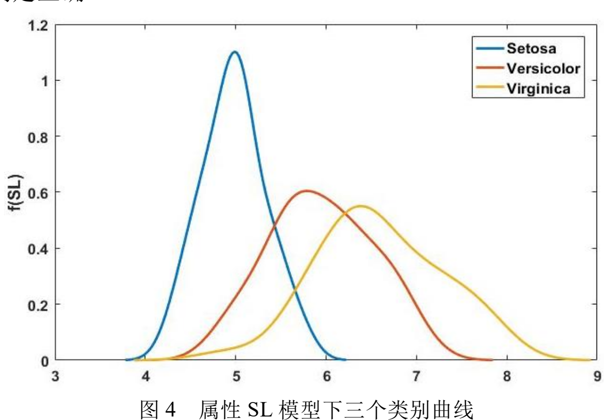


图 4　属性 SL 模型下三个类别曲线

Fig. 4　Kde curves of three classes for attribute SL

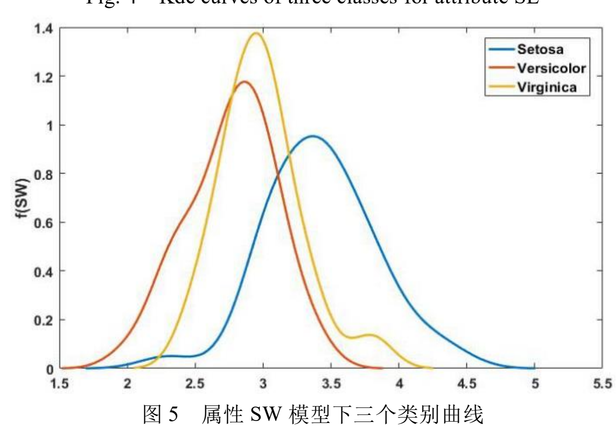


图 5　属性 SW 模型下三个类别曲线

Fig. 5　Kde curves of three classes for attribute SW

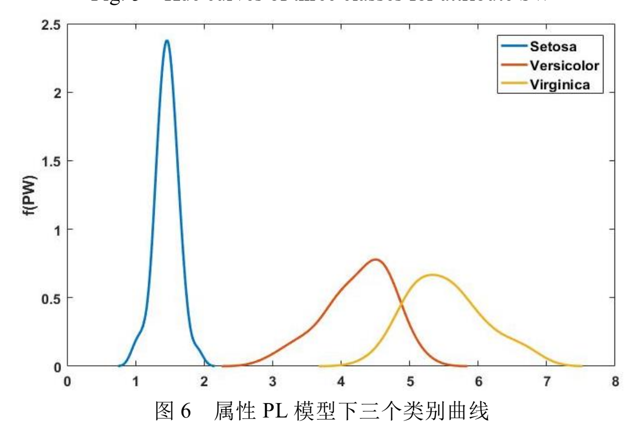


图 6　属性 PL 模型下三个类别曲线

Fig. 6　Kde curves of three classes for attribute PL

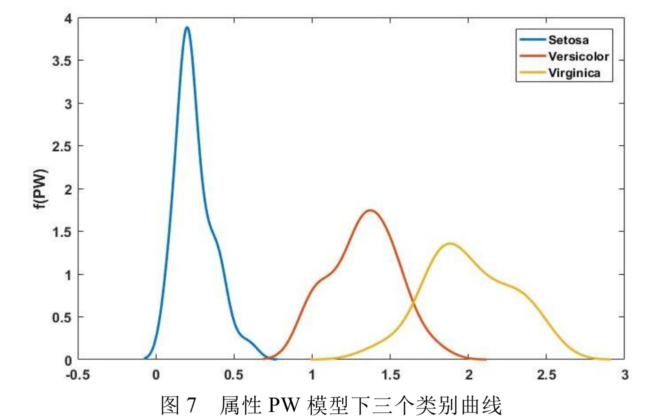


图 7　属性 PW 模型下三个类别曲线

Fig. 7　Kde curves of three classes for attribute PW

表 1　各属性对应类别的窗宽

Table 1　Bandwidth for all classes of each attribute

属性	S	C	V
SL	0.18	0.28	0.33
SW	0.21	0.17	0.17
PL	0.09	0.26	0.28
PW	0.05	0.11	0.14

表 2　各属性下类别的中心点

Table 2　Center for all classes of each attribute

属性	S	C	V
SL	5.005	5.988	6.590
SW	3.420	2.778	2.980
PL	1.447	4.265	5.523
PW	0.243	1.335	2.013

表 3　Tri-D 值

Table 3　Tri-D values

属性	S	C	V
SL	0.014	0.692	0.305
SW	0.226	1.555	0.96
PL	0	2.692	0.128
PW	0	2.271	0.078

表 4　测试样例的四条 BPA

Table 4　Four bpas of test sample

属性	BPA
SW	m({C})=0.684，m({C,V})=0.302，m({C,V,S})=0.014
SL	m({C})=0.567，m({C,V})=0.351，m({C,V,S})=0.082
PW	m({C})=0.955，m({C,V})=0.045，m({C,V,S})=0
PL	m({C})=0.967，m({C,V})=0.033，m({C,V,S})=0

3.2　核函数和窗宽选择的实验

kde 中有两个重要因素窗宽 h 和核函数，每个因素都会对结果产生影响，本节给出这两个因素变化时本文方法在 UCI 数据集上 5 折交叉验证的准确率对比。

首先是对采用不同核函数时 UCI 数据集的准确率对比由图 8 给出。总体效果来看使用高斯核函数（normal）的效果是最好的，就数据集分别看，在 Iris、Heart、Australian 三个数据集上使用不同核函数的效果相当，在 Zoo 和 Sonar 数据集上略有差距。接着在高斯核函数的条件下验证不同窗宽 h 对准确率的影响，结果由图 9 给出。可以看出采用最优化窗宽的准确率明显比使用默认窗宽要高，这是因为优化的窗宽是根据不同属性及不同类别的数据特征进行具体求解的，所以相比所有数据都用一个窗宽来讲可以更加准确地反映数据的特征，进而产生更高的准确率。

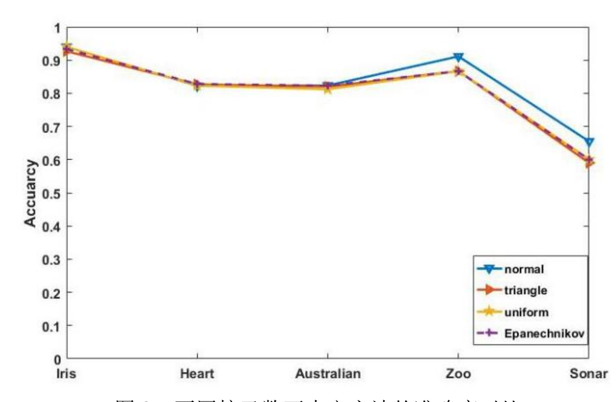


图 8　不同核函数下本文方法的准确率对比

Fig. 8　Accuracy of proposed method under different kernel functions

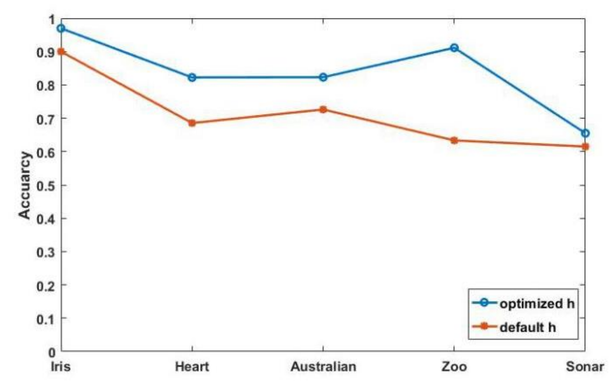


图 9　各数据集使用优化 h、默认 h 的准确率对比

Fig. 9　Accuracy of optimized h and default h on data sets

3.3　UCI 数据集上的分类实验

本实验通过与七个被大家熟知的分类器及另外三种 BPA 生成的方法进行对比实验说明本文方法的有效性。由表 5 可以看出，所采用的数据集在表中给出的样例数、类别数、属性及是否有缺失值四个方面都不尽相同，用于实验可以较全面地展现方法的有效性。5 折交叉验证的实验结果如表 6 所示。

表 5　UCI 数据集的基本信息

Table 5　Basic information of five UCI data sets

数据集	样例数	类别数	属性	是否有缺失值
Iris	150	3	4	无
Heart	270	2	13	无
Australian	690	2	14	有
Zoo	101	7	16	无
Sonar	208	2	60	无

表 6　数据集分类正确率/%

Table 6　Classification accuracy of five data sets/%

数据集	分类器							BPA 生成方法			
	NB	IB1	REPTree	SVM	SVM-RBF	MP	RBFN	方法 1[12]	方法 2[19]	方法 3[18]	本文方法
Iris	94.67	94	92	94.67	92.7	93.33	92.67	95.33	94	94.67	**96**
Heart	82.59	57.78	70.74	83.70	82.96	75.19	81.85	76.3	75.1	81.5	**83.73**
Australian	79.56	67.4	80.59	80.29	79.86	**82.32**	82.61	78.41	80.6	80.01	**82.32**
Zoo	93.1	94.1	84.2	84.1	72.3	93.1	93.1	90.5	84.1	93.1	**94.2**
Sonar	66.28	**79.78**	70.70	66.32	65.90	66.86	66.84	65.5	66.58	65.57	67.5
Average	83.24	78.61	79.65	81.82	78.74	82.16	83.41	81.21	80.08	82.97	**84.75**

通过表 6 结果本文可以看出，本文方法在各个特征都不尽相同的 UCI 数据集上的分类准确率整体优于其他几种常见的分类器，同时也优于其他几种 BPA 生成方法，其他几种 BPA 生成方法与分类器相比表现较平均。本文及对比 BPA 生成法在 Sonar 数据集上的表现均偏差，就本文方法初步考虑可能是由于特征较多，在将测试数据对应生成的多条 BPA 进行融合的过程中融合结果出现偏差导致，但整体来说，本文方法效果还是不错，说明本文方法可以有效地由包含不同特征的数据信息生成BPA即证据供D-S证据理论可以直接融合使用。

4　结束语

在 D-S 证据理论应用的过程中，基本概率指派(BPA)的生成方法仍然是一个开放性的问题，本文提出了一种基于核密度估计的 BPA 生成方法：首先构建训练数据优化的核密度属性模型；然后根据所提出的 *Tri-D* 的概念计算并分配得到测试数据的 BPA，通过 D-S 组合得出最后的结果:本文方法具有主观性弱，BPA 之间冲突小的优点。实验结果表明了所提方法的有效性；但在维数较高的数据集上进行 BPA 生成及后续应用还存在一定不足，下一步将针对此方面考虑结合改进融合方法进行思考和改进。

参考文献：

[1] Dempster A P. Upper and lower probabilities induced by a multivalued mapping [J]. Annals of Mathematical Statistics, 1967, 38 (2): 325-339.

[2] Shafer G. A mathematical theory of evidence [M]. NewJersey: Princeton University Press, 1976.

[3] 何晶晶, 蔡德胜, 介飞, 等. 利用 D-S 证据理论进行特征融合的同义实体识别 [J]. 计算机应用研究, 2018, 35 (5): 1429-1433. (He Jingjing, Cai Desheng, Jie Fei, *et al.* Synonymous entity recognition based on D-S evidence theory for feature fusion [J]. Application Research of Computers, 2018, 35 (5): 1429-1433.)

[4] 门志远, 李林宏, 张耀辉. 基于改进证据理论的齿轮技术状态评估方法 [J]. 火力与指挥控制, 2018, 43 (5): 65-68. (Men Zhiyuan, Li Linhong, Zhang Yaohui. State evaluation method of gear technology based on improved evidence theory [J]. Fire Control & Command Control, 2018, 43 (5): 65-68.)

[5] 余思奇, 景博, 黄以锋. 基于 D-S 证据理论的测试性综合评估方法 [J]. 计算机应用研究, 2014, 31 (7): 2071-2073. (Yu Siqi, Jing Bo, Huang Yifeng. Test-based comprehensive evaluation method based on D-S evidence theory [J]. Application Research of Computers, 2014, 31 (7): 2071-2073.)

[6] Xu Xiaobin, Zheng Jin, Yang Jianbo, *et al.* Data classification using evidence reasoning rule [J]. Knowledge-Based Systems, 2017, 116 (1): 144-151.

[7] 李英. 基于多分类器 DS 证据理论融合的杂草识别 [J]. 电脑编程技巧与维护, 2018 (2): 145-146, 153. (Li Ying. Weed identification based on multi-classifier DS evidence theory fusion [J]. Computer Programming Skills & Maintenance, 2018 (2): 145-146, 153.)

[8] Boccaletti S, Bianconi G, Criado R, *et al.* The structure and dynamics of multilayer networks [J]. Physics Reports, 2014, 544 (1): 1-122.

[9] Burkov A, Chaib-Draa B. Computing equilibria in discounted dynamic games [J]. Applied Mathematics & Computation, 2015, 269 (10): 863-884.

[10] Wang Zhen, Wang Lin, Szolnoki A, *et al.* Evolutionary games on multilayer networks: a colloquium [J]. European Physical Journal B, 2015, 88 (5): 124.

[11] Yager R R. A class of fuzzy measures generated from a Dempster-Shafer belief structure [J]. International Journal of Intelligent

Systems, 1999, 14 (12): 1239–1247.

[12] 蒋雯，陈运东，汤潮，等. 基于样本差异度的基本概率指派生成方法 [J]. 控制与决策, 2015, 30 (1): 71-75. (Jiang Wen, Chen Yundong, Tang Chao, *et al.* Basic probability assignment generation method based on sample difference degree [J]. Control and Decision, 2015, 30 (1): 71-75.)

[13] 涂世杰，陈航，冯刚. 证据理论与模糊函数相结合的弹道目标识别 [J]. 计算机工程与应用, 2017, 53 (6): 169-173. (Tu Shijie, Chen Hang, Feng Gang. Ballistic target recognition based on evidence theory and fuzzy function [J]. Computer Engineering and Applications, 2017, 53 (6): 169-173.)

[14] Jiang Wen, Yang Yan, Luo Yu, *et al.* Determining basic probability assignment based on the improved similarity measures of generalized fuzzy numbers [J]. International Journal of Computers Communications & Control, 2015, 10 (3): 333-347.

[15] Liu Yuting, Pal N R, Marathe A R, *et al.* Weighted fuzzy Dempster-Shafer framework for multimodal information integration [J]. IEEE Trans on Fuzzy Systems, 2018, 26 (1): 338-352.

[16] Zhang Chenwei, Hu Yong, Chan Felix T. S, *et al.* A new method to determine basic probability assignment using core samples [J]. Knowledge-Based Systems, 2014, 69 (1): 140-149.

[17] Deng Xinyang, Liu Qi, Deng Yong, *et al.* An improved method to construct basic probability assignment based on the confusion matrix for classification problem [J]. Information Sciences, 2016, 340-341: 250-261.

[18] 高智勇，董荣光，高建民，等. 采用聚类特征的基本概率分配生成方法及应用 [J]. 西安交通大学学报, 2016, 50 (10): 8-14. (Gao Zhiyong, Dong Rongguang, Gao Jianmin, *et al.* Basic probability distribution generation method and application using clustering features [J]. Journal of Xi'an Jiaotong University, 2016, 50 (10): 8-14.)

[19] 陈亚男，任圣君. 大样本条件下证据基本概率指派函数生成方法研究 [J]. 信息工程大学学报, 2017, 18 (4): 118-122. (Chen Yanan, Ren Shengjun. Research on generating probability assignment function of evidence under large sample conditions [J]. Journal of Information Engineering University, 2017, 18 (4): 118-122.)

[20] Mahadevan S, Deng Yong, Xu Peida, *et al.* A new method to determine basic probability assignment from training data [J]. Knowledge-Based Systems, 2013, 46 (1): 69-80.

[21] 李民政，蓝剑平. 改进的基于 Pignistic 概率距离的证据组合方法 [J]. 桂林电子科技大学学报，2017, 37 (4): 63-67. (Li Minzheng, Lan Jianping. Improved evidence combination method based on Pignistic probability distance [J]. Journal of Guilin University of Electronic Technology, 2017, 37 (4): 63-67.)

[22] Rosenblatt M. Remarks on some nonparametric estimates of a density function [J]. Annals of Mathematical Statistics, 1956, 27 (3): 832-837.

[23] Parzen E. On estimation of a probability density function and mode [J]. Annals of Mathematical Statistics, 1962, 33 (3): 1065-1076.

[24] Bolancé C, Guillen M, Nielsen J P. Kernel density estimation of actuarial loss functions [J]. Insurance Mathematics & Economics, 2009, 32 (1): 19-36.

[25] Arthur D. K-means+: the advantages of careful seeding [C]// Proc of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2007: 1027-1035.